

Flow Network based Generative Models for Non-Iterative Diverse Candidate Generation

Year: 2021 | Citations: 419 | Authors: Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, Y. Bengio

Abstract

This paper is about the problem of learning a stochastic policy for generating an object (like a molecular graph) from a sequence of actions, such that the probability of generating an object is proportional to a given positive reward for that object. Whereas standard return maximization tends to converge to a single return-maximizing sequence, there are cases where we would like to sample a diverse set of high-return solutions. These arise, for example, in black-box function optimization when few rounds are possible, each with large batches of queries, where the batches should be diverse, e.g., in the design of new molecules. One can also see this as a problem of approximately converting an energy function to a generative distribution. While MCMC methods can achieve that, they are expensive and generally only perform local exploration. Instead, training a generative policy amortizes the cost of search during training and yields to fast generation. Using insights from Temporal Difference learning, we propose GFlowNet, based on a view of the generative process as a flow network, making it possible to handle the tricky case where different trajectories can yield the same final state, e.g., there are many ways to sequentially add atoms to generate some molecular graph. We cast the set of trajectories as a flow and convert the flow consistency equations into a learning objective, akin to the casting of the Bellman equations into Temporal Difference methods. We prove that any global minimum of the proposed objectives yields a policy which samples from the desired distribution, and demonstrate the improved performance and diversity of GFlowNet on a simple domain where there are many modes to the reward function, and on a molecule synthesis task.